

Article

Bayesian Compressive Sensing Based Optimized Node Selection Scheme in Underwater Sensor Networks

Ruisong Wang ¹ , Gongliang Liu ¹, Wenjing Kang ^{1,*}, Bo Li ¹, Ruofei Ma ¹ and Chunsheng Zhu ²

¹ Harbin Institute of Technology, Weihai 264209, China; mathwrs@163.com (R.W.); liugl@hit.edu.cn (G.L.); libo1983@hit.edu.cn (B.L.); maruofei@hit.edu.cn (R.M.)

² Department of Electrical and Computer Engineering, The University of British Columbia, Vancouver, BC V6T 1Z4, Canada; chunsheng.tom.zhu@gmail.com

* Correspondence: kwjq@hit.edu.cn; Tel.: +86-158-6630-0767

Received: 20 May 2018; Accepted: 31 July 2018; Published: 6 August 2018



Abstract: Information acquisition in underwater sensor networks is usually limited by energy and bandwidth. Fortunately, the received signal can be represented sparsely on some basis. Therefore, a compressed sensing method can be used to collect the information by selecting a subset of the total sensor nodes. The conventional compressed sensing scheme is to select some sensor nodes randomly. The network lifetime and the correlation of sensor nodes are not considered. Therefore, it is significant to adjust the sensor node selection scheme according to these factors for the superior performance. In this paper, an optimized sensor node selection scheme is given based on Bayesian estimation theory. The advantage of Bayesian estimation is to give the closed-form expression of posterior density function and error covariance matrix. The proposed optimization problem first aims at minimizing the mean square error (MSE) of Bayesian estimation based on a given error covariance matrix. Then, the non-convex optimization problem is transformed as a convex semidefinite programming problem by relaxing the constraints. Finally, the residual energy of each sensor node is taken into account as a constraint in the optimization problem. Simulation results demonstrate that the proposed scheme has better performance than a conventional compressed sensing scheme.

Keywords: Bayesian estimation; compressed sensing; network lifetime; underwater sensor network

1. Introduction

Compared to the traditional sensor networks [1,2], the underwater sensor network is more restricted to the resource. The study of underwater sensor networks has become a hot topic and has attracted more attention in recent years [3–6]. For example, routing protocols and secure communication for underwater sensor networks are investigated in [7,8], respectively. However, none of them focus on the full use of node energy resources. However, the energy of each sensor node is fixed and limited in underwater sensor networks. A traditional information acquisition scheme needs to collect the information of all the sensor nodes so that it costs too much energy. Hence, some works are provided for decreasing the price of sensor networks [9,10]. In addition, it is difficult for the sink node to distinguish the signals from different sensor nodes. An efficient and affordable scheme needs to be proposed for superior performance. A compressed sensing method can be used to solve it as long as the original signal is sparse [11]. In the underwater sensor network, many kinds of signals received by sink nodes can be represented sparsely on some basis such as temperature and salinity. Therefore, we can collect a part of the total sensor nodes and then reconstruct the original signal. The compressed sensing scheme can not only decrease the cost but also improve the efficiency. The state of network and the correlation of nodes are not considered due to the randomness of node

selection in a conventional compressed sensing scheme. Hence, the performance of the compressed sensing scheme can be further improved by considering them.

1.1. Related Works and Motivation

Compressed sensing, which has been studied by many scholars [12–15], is a vital technique for decreasing the samples in the information acquisition. For example, the authors in [16] applied a compressed sensing method to multipath channel estimation. In [17], the authors designed an optimized compressed sensing matrix and applied it to the sparse channel estimation. In [18], the directions of arrival estimation were studied by a compressed sensing method and it is robust and very accurate. Compressed sensing theory is also used to develop the practical incoherent undersampling schemes for rapid MR imaging [19]. By using a compressed sensing method, a semi-analytic scheme is developed to solve the iterative clipping noise recovery in [20]. By taking advantage of the compressed sensing method, the authors in [21] addressed nonlinear inverse scattering problems. Furthermore, examples with experimental data also showed the effectiveness of the compressed sensing method in position and shape reconstruction [22]. In addition, compressed sensing theory can be used to cancer imaging about the induced magnetic anomaly [23]. Therefore, it is of significance to develop the application of the compressed sensing technique. In recent years, Bayesian compressive sensing has been widely investigated due to its advantages of giving posterior density function and error covariance matrix. For instance, a Bayesian compressive sensing method is used to estimate complex-valued targets in wireless localization systems [24]. Thermal imaging is studied with an application of Bayesian compressive sensing in [25]. In [26], the authors provided an estimation for direction of arrival based on Bayesian compressive sensing and gave a Kalman filter to analyze the signals. A Bayesian compressive sensing framework was used to develop the channel state information (CSI) acquisition scheme by considering the parametric sparsity and channel model in [27]. In [28], the authors studied Bayesian framework and provided some useful results and algorithms. In [29], ultra-wideband channel estimation is investigated by making use of Bayesian estimation theory. The authors also analyzed the performance and the application situations of this scheme.

Although there are some works about the application of compressed sensing in underwater sensor networks [30–34], few studies focused on the effect of the sensor node selection scheme. For underwater sensor networks, the energy is strictly limited because the underwater sensor node can not be charged again. It is significant to consider the sustainability of the network when the sink node determines the selection set of sensor nodes. In addition, to obtain more accurate reconstruction results, the performance of reconstruction is worthy of studying. Motivated by the analysis above, we aim at providing a novel sensor node selection scheme for improving the reconstruction performance and the network lifetime. By making full use of the advantage of Bayesian theory, the mean square error (MSE) of estimation can be utilized to optimize the sensor node selection scheme.

1.2. Contributions

The paper considers an underwater sensor network where there is a sink node and the sensor nodes transmit the information to it directly. Suppose that the sink node knows the location of each sensor node. Combining Bayesian theory and convex optimization theory, we investigated an optimized sensor node selection scheme based on compressed sensing in underwater sensor networks. The proposed scheme can be obtained by solving an optimization problem. Both the network lifetime and reconstruction accuracy performance are considered in the optimization problem. In this paper, our main contributions can be summarized as follows:

- The paper focuses on the improvement of performance of a compressed sensing scheme. First, a Bayesian estimation theory is provided for the signal reconstruction. Then, based on Bayesian estimation theory, the closed-form expression of posterior density function and error covariance matrix are given. By a maximum posteriori estimation, the noise variance can be obtained by updating iteratively.

- The sensor node selection scheme is transformed as a sensing matrix design problem. By using the error covariance matrix of Bayesian estimation and regarding the sensing matrix as a variable, the sensing matrix design problem can be treated as an optimization problem of minimizing the MSE. Because the proposed optimization problem is non-convex, the optimal solution is difficult to obtain. By relaxing the integer constraint as a continuous constraint, the proposed optimization problem becomes a convex semidefinite programming problem that can be solved efficiently.
- The sustainability of networks is considered in the proposed optimization problem. To prolong the network lifetime, the scheme aims at selecting the sensor nodes holding more residual energy. This idea is transformed into a constraint in the optimization problem.

The proposed sensor node selection scheme is evaluated by simulation results. Simulation results demonstrate that the optimized node selection scheme has superior performance compared to the traditional one.

1.3. Paper Organization

The rest of the paper is organized as follows. The compressed sensing theory, Bayesian method, and the underwater channel model are introduced in Section 2. We propose the optimization problem about the MSE and the energy limit in Section 3. Section 4 provides the simulation results to show the superior performance of the proposed scheme. A conclusion of the paper is given in Section 5.

1.4. Notation

The following notations are used in the rest of the paper. Let the bold fonts denote vectors or matrices. The transpose operations are denoted by $(\cdot)^T$, respectively. Let $\|\cdot\|_2$ denote the (Euclidean) vector norms. If there are no special instructions, the inequality between two vectors refers to the element of the vectors. The inverse of matrix is denoted as $(\cdot)^{-1}$. \mathbf{I}_N denotes the identity matrix with the dimension $N \times N$.

2. System Model and Problem Formulation

2.1. System Model

The paper considers an underwater sensor network where a sink node is located in the center of a circular region and N sensor nodes are uniformly distributed in this range. The radius is r . Each sensor collects information such as temperature, salinity, and so on and then transmits them to a sink node. In the traditional information acquisition scheme, the sink node collects the information of all the sensor nodes that consume too much energy. To overcome this difficulty, a compressed sensing method can be applied to reduce the number of measurements. Suppose that the total signal $\mathbf{x} = [x(1), x(2), \dots, x(N)]^T$, where $i = 1, 2, \dots, N$ can be expressed as

$$\mathbf{x} = \sum_{i=1}^N s(i) \Psi(i) = \Psi \mathbf{s}, \quad (1)$$

where $\Psi = [\Psi(1), \Psi(2), \dots, \Psi(N)]$ is a basis matrix and \mathbf{s} is a sparse vector that there are $K \ll N$ nonzero elements and other elements are zero. The K is called the sparsity of vector \mathbf{s} . The vector \mathbf{x} is called compressible if the conditions above are satisfied.

According to the compressed sensing theory, the sink node can select $M \ll N$ measurements to reconstruct the original signal only if it is compressible in the basis Ψ . Let $\Phi \in \mathbb{R}^{M \times N}$ denote the sensor node selection matrix which is independent from Ψ . The structure of node selection matrix Φ can be represented as follows:

$$\begin{cases} \sum_{i=1}^M \Phi_{ij} \leq 1, j = 1, 2, \dots, N, \\ \sum_{j=1}^N \Phi_{ij} = 1, i = 1, 2, \dots, M, \\ \Phi_{ij} \in \{0, 1\}, i = 1, 2, \dots, M, j = 1, 2, \dots, N. \end{cases} \quad (2)$$

If the i th element in the j th row of node selection matrix Φ is 1, it means that the i th node is selected. Then, the measurement vector can be obtained

$$\mathbf{y} = \Phi \Psi \mathbf{s} + \mathbf{n} = \mathbf{A} \mathbf{s} + \mathbf{n}, \quad (3)$$

where $\mathbf{A} = \Phi \Psi$ is the projection matrix. The vector \mathbf{n} is independent identically distributed (i.i.d) additive white Gaussian noise with mean zero and variance σ^2 :

$$p(\mathbf{n}) = \prod_{i=1}^N \mathcal{N}(n_i | 0, \sigma^2). \quad (4)$$

The paper aims at reconstructing the sparse vector \mathbf{s} instead of reconstructing the original signal vector \mathbf{x} directly. From Equation (3), there are N variables and $M \ll N$ equalities. It is impossible to obtain the unique solution. However, notice that the vector \mathbf{s} is sparse and has $K \ll N$ nonzero elements, and it makes the reconstruction possible if the projection matrix \mathbf{A} meets the restricted isometry property (RIP). The reconstructed solution $\hat{\mathbf{s}}$ can be obtained from the optimization problem as follows:

$$\begin{aligned} & \arg \min \|\mathbf{s}\|_0, \\ & s.t. \|\mathbf{y} - \mathbf{A} \mathbf{s}\|_2 \leq \varepsilon. \end{aligned} \quad (5)$$

However, the optimization problem above is non-deterministic polynomial (NP)-hard due to the non-convexity of the objective function. To relax the original problem as a convex problem, a convex $\|\cdot\|_1$ is usually used to replace the $\|\cdot\|_0$. Then, the transformed problem can be represented as follows:

$$\begin{aligned} & \arg \min \|\mathbf{s}\|_1, \\ & s.t. \|\mathbf{y} - \mathbf{A} \mathbf{s}\|_2 \leq \varepsilon. \end{aligned} \quad (6)$$

The problem above can be solved effectively by the CVX tool cabinet. The original signal $\hat{\mathbf{x}} = \Psi \hat{\mathbf{s}}$ can be recovered by using the solution $\hat{\mathbf{s}}$. In the underwater information sampling, the signal is usually sparse due to the spatial correlation. For example, the temperature signal is sparse in the frequency domain, i.e., it can be expressed sparsely by Fourier basis. Therefore, it is significant to apply a compressed sensing technique for underwater data sampling.

2.2. Bayesian Compressive Sensing

In this subsection, the sparse vector \mathbf{s} is reconstructed by measurement \mathbf{y} according to Bayesian estimation. Compared with the traditional compressive sensing reconstruction algorithm, the advantage of Bayesian compressive sensing is that the noise variance can also be estimated. In addition, the posterior probability density function can be given through some prior information. The main idea of Bayesian estimation is taking full advantage of reasonably assumed prior information and the prior information including the distribution of noise and the measurement \mathbf{y} . Firstly, the Gaussian likelihood model of the measurement \mathbf{y} can be written as follows [28]:

$$p(\mathbf{y} | \mathbf{s}, \sigma^2) = (2\pi\sigma^2)^{-\frac{M}{2}} e^{-\frac{\|\mathbf{y} - \mathbf{A} \mathbf{s}\|_2^2}{2\sigma^2}}. \quad (7)$$

The posterior of the sparse vector \mathbf{s} needs to be given to estimate the final solution. According to the Bayes' rule, the prior information of the sparse vector \mathbf{s} needs to be known. In general, Laplace prior is used in many works. However, it is not conjugate with the Gaussian likelihood model. The closed form solution of posterior distribution is difficult to obtain. Hence, a hierarchical prior is used in this paper instead of using it. Suppose that the element of the sparse vector \mathbf{s} is an independent Gaussian distribution with zero mean and variance α_i^{-1} . Then, the distribution of the sparse vector \mathbf{s} can be shown:

$$\begin{aligned}
 p(\mathbf{s}|\boldsymbol{\alpha}) &= \prod_{i=1}^N \mathcal{N}(s_i|0, \alpha_i^{-1}) \\
 &= \prod_{i=1}^N (2\pi)^{-\frac{1}{2}} \alpha_i^{\frac{1}{2}} e^{-\frac{\alpha_i s_i^2}{2}},
 \end{aligned} \tag{8}$$

where $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_N)^T$ is the hyperparameter.

Hierarchical priors about the hyperparameters $\boldsymbol{\alpha}$, σ^2 and the sparse vector \mathbf{s} are given in Appendix A. It is conjugate with Gaussian distribution. The marginal likelihood function $p(\mathbf{y}|\boldsymbol{\alpha}, \sigma^2)$ can be computed by integrating $p(\mathbf{y}|\mathbf{s}, \sigma^2)p(\mathbf{s}|\boldsymbol{\alpha})$

$$\begin{aligned}
 p(\mathbf{y}|\boldsymbol{\alpha}, \sigma^2) &= \int p(\mathbf{y}|\mathbf{s}, \sigma^2)p(\mathbf{s}|\boldsymbol{\alpha})d\mathbf{s} \\
 &= (2\pi)^{-\frac{M}{2}} |\mathbf{C}|^{-\frac{1}{2}} e^{-\frac{\mathbf{y}^T \mathbf{C}^{-1} \mathbf{y}}{2}},
 \end{aligned} \tag{9}$$

where $\mathbf{C} = \sigma^2 \mathbf{I} + \boldsymbol{\Lambda} \mathbf{A}^{-1} \mathbf{A}^T$ and $\boldsymbol{\Lambda}$ is a diagonal matrix of the hyperparameters $\boldsymbol{\alpha}$.

According to the assumption of the hierarchical prior and the Gaussian likelihood model, the closed form solution of the posterior distribution can be derived:

$$\begin{aligned}
 p(\mathbf{s}|\boldsymbol{\alpha}, \sigma^2, \mathbf{y}), \\
 = (2\pi)^{-\frac{N}{2}} |\boldsymbol{\Sigma}|^{-\frac{1}{2}} e^{-\frac{(\mathbf{s}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{s}-\boldsymbol{\mu})}{2}},
 \end{aligned} \tag{10}$$

where the mean and covariance matrix are $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ as follows:

$$\begin{aligned}
 \boldsymbol{\mu} &= \sigma^{-2} \boldsymbol{\Sigma} \mathbf{A}^T \mathbf{y} \\
 \boldsymbol{\Sigma} &= (\sigma^{-2} \mathbf{A}^T \mathbf{A} + \boldsymbol{\Lambda})^{-1}.
 \end{aligned} \tag{11}$$

It can be seen that $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are functions of the hyperparameters $\boldsymbol{\alpha}$ and σ^2 which are unknown. It is necessary to estimate them to obtain the mean and variance of the sparse vector \mathbf{s} . The hyperparameters' estimation can be obtained by maximizing $p(\mathbf{y}|\boldsymbol{\alpha}, \sigma^2)p(\boldsymbol{\alpha})p(\sigma^{-2})$. Due to a uniform distribution assumption of $\boldsymbol{\alpha}$ and σ^2 , the hyperparameters' estimation can be obtained by maximizing the marginal likelihood function $p(\mathbf{y}|\boldsymbol{\alpha}, \sigma^2)$ directly, or equivalently maximizing the logarithm of $p(\mathbf{y}|\boldsymbol{\alpha} : \sigma^2)$.

$$\begin{aligned}
 \mathcal{L} &= \log p(\mathbf{y}|\boldsymbol{\alpha}, \sigma^2) \\
 &= -\frac{1}{2} (M \log(2\pi) + \log |\mathbf{C}| + \mathbf{y}^T \mathbf{C}^{-1} \mathbf{y}).
 \end{aligned} \tag{12}$$

By taking the derivative of \mathcal{L} about $\boldsymbol{\alpha}$ and setting it as zero, the estimation of $\boldsymbol{\alpha}$ can be given as follows:

$$\alpha_i = \frac{1}{\mu_i^2 + \Sigma_{ii}}, \tag{13}$$

where Σ_{ii} is the i th element of the covariance matrix $\boldsymbol{\Sigma}$. By using the same method, the estimation of σ^2 can be obtained as follows:

$$\sigma^2 = \frac{\|\mathbf{y} - \mathbf{A}\boldsymbol{\mu}\|_2^2}{M - N + \sum_{i=1}^N \alpha_i \Sigma_{ii}}. \tag{14}$$

The derivations of Equations (13) and (14) are similar to [35]. It can be seen that the hyperparameters $\boldsymbol{\alpha}$ and σ^2 are functions of the mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. The expectation maximization (EM) algorithm shown as Algorithm 1 can be used to update them iteratively. The iteration is terminated when there is a convergency value. Finally, the mean $\boldsymbol{\mu}$ is regarded as the estimation of the sparse vector \mathbf{s} . The original signal \mathbf{x} can be reconstructed by using $\boldsymbol{\Psi}\boldsymbol{\mu}$.

Algorithm 1 Expectation maximization algorithm for Bayesian estimation [35].

- 1: Initialize $\alpha_i[0], \forall i = 1, 2, \dots, N$ and $\sigma^2[0]$, set tolerance as ε and maximum number of iterations as K .
 - 2: **for** $k = 1 : K$ **do**
 - 3: Calculate the mean $\boldsymbol{\mu}[k]$ and covariance matrix $\boldsymbol{\Sigma}[k]$.
 - 4: Update the hyperparameters $\alpha_i[k], \forall i = 1, 2, \dots, N$ and $\sigma^2[k]$ according to the results of step 3.
 - 5: **if** $|\alpha_i[k] - \alpha_i[k - 1]| < \varepsilon$ **then**
 - 6: Break.
 - 7: **end if**
 - 8: **end for**
 - 9: Output the final mean $\boldsymbol{\mu}[k]$ and covariance matrix $\boldsymbol{\Sigma}[k]$ and reconstruct the signal as $\boldsymbol{\Psi}\boldsymbol{\mu}$.
-

2.3. Underwater Channel Model

In the underwater environment, the signal transmission is usually achieved by acoustic transmission, which has characteristics of low speed and has long delay. Due to the reflection, refraction, and current, the pathloss in underwater transmission is large. The pathloss mode can be given as follows:

$$A(d_s, f) = d_s^\tau \phi(f)^{d_s}, \quad (15)$$

where d_s is the distance between the sensor node and sink node, f is the carrier frequency, and τ is a constant which is usually set as 1.5.

The term $\phi(f)$ is called as absorption coefficient and given as follows:

$$\phi(f) = 10^{\frac{\lambda(f)}{10}}. \quad (16)$$

The $\lambda(f)$ can be estimated by Thorps formula as

$$\lambda(f) = \frac{0.11f^2}{1 + f^2} + \frac{44f^2}{4400 + f^2} + \frac{2.75}{10^4}f^2 + 0.003. \quad (17)$$

Ideal power control is assumed in this paper. Then, suppose that the power required by the sink node is P_0 , and the sensor nodes need to take the power as least $P_0/A(d, f)$ to transmit the signal. Based on the analysis above, the energy cost model can be provided as follows:

$$E_i = \begin{cases} E_c + E_u A^{-1}(d_s, f), & \text{if the node is selected,} \\ E_c, & \text{otherwise.} \end{cases} \quad \forall i = 1, 2, \dots, N, \quad (18)$$

where E_c denotes the circuit energy consumption and E_u is the energy cost that can meet the requirement of minimum power without path loss. Compared with $E_u A^{-1}(d_s, f)$, the circuit energy consumption can be ignored because it is the same for each sensor node. The main focus of this paper is that the node selection contributes to the influence of network lifetime due to different energy consumption of each node.

3. Optimized Node Selection Scheme

In the traditional sensor nodes' selection scheme, each sensor node is selected by the same probability, which doesn't consider the accuracy of reconstruction results and only guarantees that the signal can be reconstructed. Based on the Bayesian estimation scheme, an optimized sensor nodes' selection scheme can be driven by considering the corresponding MSE and the network lifetime.

3.1. Scheme for Minimizing MSE of Estimation

Let $\tilde{\Phi} = \Phi^T \Phi$ be a diagonal matrix and $\tilde{\Phi} \in \mathbb{R}^{N \times N}$ denote the sensor nodes' selection matrix. If the i th diagonal element is 1, it means that the i th sensor node is selected. If the i th diagonal element is 0, it means that the i th sensor node is not selected. It is well known that the MSE of Bayesian estimation is the trace of error covariance matrix. By setting the sensor nodes selection matrix $\tilde{\Phi}$ as a variable and minimizing the MSE, the following optimization problem can be obtained:

$$\begin{aligned} \min_{\tilde{\Phi}} \quad & \text{Trace}(\sigma^{-2} \Psi^T \tilde{\Phi} \Psi + \Lambda)^{-1}, \\ \text{s.t.} \quad & \tilde{\Phi}_{ii} \in \{0, 1\}, i = 1, 2, \dots, N, \\ & \text{Trace}(\tilde{\Phi}) = M. \end{aligned} \quad (19)$$

The first constraint condition decides whether the node is selected or not. The second constraint condition means that the number of selected nodes is M . Notice that the proposed optimization problem is non-convex due to the non-convexity of the first constraint condition, and it is difficult to obtain the optimal solution of it. To overcome this difficulty, the constraint condition is relaxed:

$$\begin{aligned} \min_{\tilde{\Phi}} \quad & \text{Trace}(\sigma^{-2} \Psi^T \tilde{\Phi} \Psi + \Lambda)^{-1}, \\ \text{s.t.} \quad & 0 \leq \tilde{\Phi}_{ii} \leq 1, i = 1, 2, \dots, N, \\ & \text{Trace}(\tilde{\Phi}) = M. \end{aligned} \quad (20)$$

By introducing a slack variable \mathbf{Q} and applying the Schur's complement, the optimization problem can be rewritten as follows:

$$\begin{aligned} \min_{\tilde{\Phi}, \mathbf{Q}} \quad & \text{Trace}(\mathbf{Q}), \\ \text{s.t.} \quad & \begin{pmatrix} \sigma^{-2} \Psi^T \tilde{\Phi} \Psi + \Lambda & \mathbf{I}_N \\ \mathbf{I}_N & \mathbf{Q} \end{pmatrix} \succeq 0, \\ & 0 \leq \tilde{\Phi}_{ii} \leq 1, i = 1, 2, \dots, N, \\ & \text{Trace}(\tilde{\Phi}) = M. \end{aligned} \quad (21)$$

According to convex optimization theory, the optimization problem above is a semidefinite programming problem. A CVX tool cabinet in Matlab (R2017a, The MathWorks Inc., Natick, MA, USA) can be used to solve it effectively. There may be some other methods to solve it more efficiently, but it is beyond the scope of this paper.

3.2. Scheme for Prolonging the Network Lifetime

The network collects the information by putting the sensor nodes under the water. According to the feature of underwater environment, the channel differences of the sensor nodes are large. For example, the node with weak channel condition needs to take more power to transmit the signal so that the sink node can obtain the complete signal. However, the energy of each node is fixed. If there is a node with no energy prematurely, the sensor nodes' network will be invalid. This leads to a serious waste of resources. Therefore, it is significant to provide a new node selection scheme for prolonging the lifetime of the sensor network. Assume that the initial energy of each sensor node is E_0 . Let E_{ij} denote the i th sensor node's energy in j th selection. Then, let $P_{ij} = \frac{E_{ij}}{\sum_{i=1}^N E_{ij}}$ denote the proportion of the i th sensor node's energy with the total energy in j th selection. The scheme aims at selecting the sensor nodes that hold more energy. Therefore, the optimization problem with the energy constraint condition in j th selection can be rewritten as follows:

$$\begin{aligned}
& \min_{\tilde{\Phi}, \mathbf{Q}} \text{Trace}(\mathbf{Q}), \\
& \text{s.t.} \begin{pmatrix} \sigma^{-2} \mathbf{\Psi}^T \tilde{\Phi} \mathbf{\Psi} + \Lambda & \mathbf{I}_N \\ \mathbf{I}_N & \mathbf{Q} \end{pmatrix} \succeq 0, \\
& \text{Trace}(\tilde{\Phi} \mathbf{P}_j) \geq C, \\
& 0 \leq \tilde{\Phi}_{ii} \leq 1, i = 1, 2, \dots, N, \\
& \text{Trace}(\tilde{\Phi}) = M,
\end{aligned} \tag{22}$$

where $\mathbf{P}_j = \text{diag}(P_{1j}, P_{2j}, \dots, P_{Nj})^T$ and C is a constant threshold. From the energy constraint condition, it can be seen that the sensor nodes with less energy are selected by a small probability. If the sensor nodes with less energy are selected, it is possible that the energy constraint condition is not satisfied because the proportion values are small. For the final solution matrix, the M sensor nodes corresponding to the largest values of diagonal element in the solution matrix are chosen. The detail can be found in Algorithm 2.

Algorithm 2 Optimized node selection algorithm based on convex relaxation.

- 1: Initialize the number of selected nodes M , threshold C , the number of selection as $j = 0$, and \mathbf{P}_0 .
 - 2: **while** The residual energy of each node is non-negative. **do**
 - 3: $j = j + 1$.
 - 4: Solve the optimization problem (22) by using CVX tool cabinet in matlab.
 - 5: The M sensor nodes corresponding to the largest values of diagonal element in the solution matrix $\tilde{\Phi}$ are chosen for data transmission.
 - 6: The signal is reconstructed by using Algorithm 1 or orthogonal matching pursuit (OMP) algorithm.
 - 7: Calculate \mathbf{P}_j according to the optimized solution matrix $\tilde{\Phi}$.
 - 8: **end while**
 - 9: Output the number of selection j and \mathbf{P}_j .
-

4. Simulation Results and Discussion

In this section, simulation results are provided for comparing our proposed optimized sensor nodes selection scheme with conventional compressed sensing scheme [31]. In [31], each sensor node sends data randomly with a probability or the fusion center deciding the nodes to send data randomly. The selection of nodes with a random scheme doesn't consider the state and correlation of each node. OMP is an efficient algorithm that is usually used to reconstruct a signal [36]. This algorithm is also compared with our proposed algorithm in the simulation. We consider a sink node in the center of a circular region with the radius $r = 500$ m. The sensor nodes are uniformly distributed in this region. For decreasing the computation complexity, the number of the sensor nodes is set as 64. The minimum distance between sink node and sensor node is 50 m. The minimum distance among sensor nodes is 40 m. Each sensor node acquires information and transmits it to sink node. The sink node collects all the information and requires that the minimum received power is $P_0 = 50$ mW. The discrete Fourier basis is applied in the evaluation. Assume that the energy of each sensor node is the same and is set as 600 J. The carrier frequency is assumed as 10 kHz. The constant C is set as $\frac{M}{N}$. The simulation shows mainly the performance of the reconstruction error and the lifetime of the sensor network. The reconstruction error is defined as follows:

$$R_e = \frac{\|\mathbf{x} - \hat{\mathbf{x}}\|_2}{\|\mathbf{x}\|_2}, \tag{23}$$

where x is the original signal and \hat{x} is the reconstructed signal. If the residual energy of one sensor node is full, the sensor network is regarded as invalid. The lifetime of the sensor network is defined as the times that the sensor network can collect the information until the first sensor node died.

In Figure 1, the performance of reconstruction error is evaluated with the number of selected sensor nodes. It can be seen that the reconstruction error is decreasing when the number of selected sensor nodes increases. The proposed optimized sensor nodes selection scheme is superior to the conventional compressed sensing scheme. The performance of reconstruction scheme by using optimized Bayesian estimation outperforms the scheme with conventional Bayesian estimation. The reconstruction scheme by using optimized Bayesian estimation is better than the OMP algorithm and the optimized OMP algorithm when the number of selected sensor nodes is more than 30. When the number of selected sensor nodes is less than 30, the optimized OMP algorithm has better performance than the OMP algorithm and the optimized Bayesian estimation scheme. It is reasonable because the prior information is little when the number of selected sensor nodes is few. It is not suitable to use a Bayesian estimation scheme.

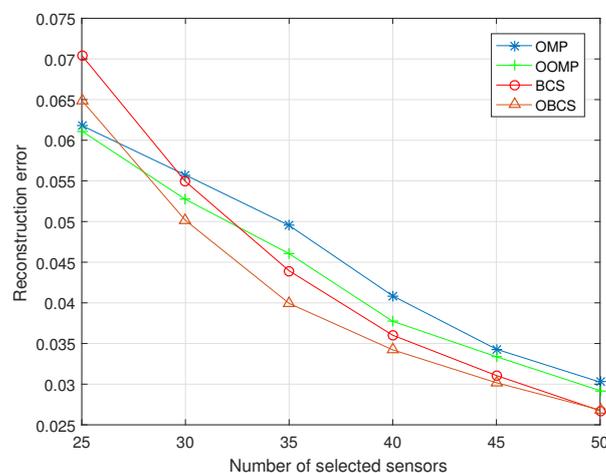


Figure 1. Reconstruction error versus different number of selected sensor nodes.

Figure 2 shows the network lifetime versus the number of selected sensor nodes. When the number of selected sensor nodes grows, the network lifetime decreases. This is because the feasible selection schemes are less as the number of selected sensor nodes grows larger. It is obvious that the optimized sensor nodes selection scheme has a significant improvement than the conventional sensor nodes' selection scheme. For example, when the number of selected sensor nodes is 25, the network lifetime with optimized sensor nodes selection scheme is 5.6 times more than conventional sensor nodes selection scheme.

This paper proposed an optimized sensor nodes selection scheme for compressed information acquisition based on Bayesian estimation. The proposed scheme is not limited to reconstructing the signal. In practice, the sink node can select different schemes to reconstruct the signal according to different performance requirements. For example, when the reconstructed signal is required to be accurate as much as possible, it is suitable to increase the number of selected sensor nodes and apply Bayesian estimation with an optimized selection scheme to reconstruct the signal. On the contrary, when the reconstruction error is not limited strictly, it is better to decrease the number of selected sensor nodes and use the OMP algorithm with an optimized selection scheme to reconstruct the signal for prolonging the network lifetime.

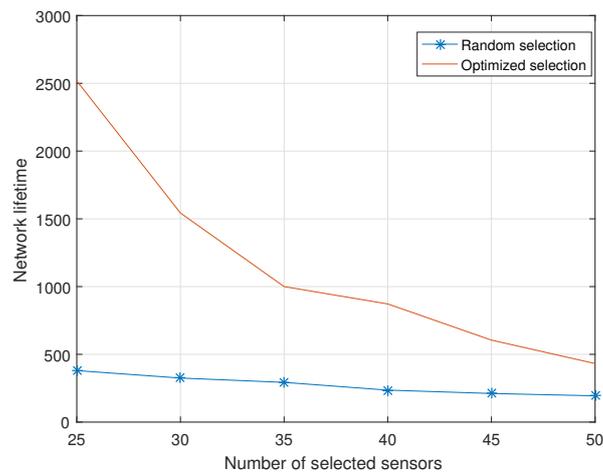


Figure 2. Network lifetime versus different number of selected sensor nodes.

Figure 3 shows the remaining energy of the sensor network versus the number of selected sensor nodes. It can be observed that the remaining energy of sensor network grows when the number of selected sensor nodes increases. We can see that the optimized sensor nodes' selection scheme has more energy efficiency than the conventional sensor nodes' selection scheme. For example, when the number of selected sensor nodes is 30, the optimized scheme makes full use of 2.3 times energy to prolong network lifetime more.

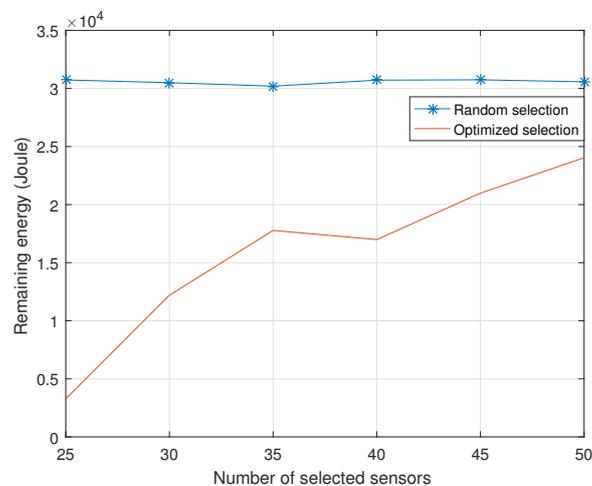


Figure 3. Remaining energy versus different number of selected sensor nodes.

In Figure 4, the remaining energy of each sensor node is provided. The number of selected sensor nodes is set as 25. It can be seen that the remaining energy of each sensor is homogeneous for the optimized sensor nodes selection scheme. However, for the conventional sensor nodes selection scheme, the remaining energy of each sensor is disorderly. The remaining energy of some sensor nodes is few, and it is adequate for many sensor nodes. Thus, the use of random node selection scheme leads to the wasting of total energy.

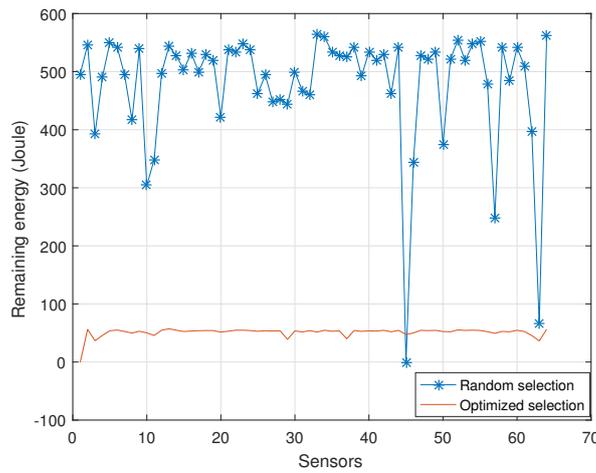


Figure 4. Remaining energy of each sensor node.

Figure 5 shows the energy efficiency versus different number of selected sensor nodes. The definition of energy efficiency is the proportion of network lifetime and the total energy cost. It represents the average time of information acquisition per unit energy. When the number of selected sensor nodes increases, the energy efficiency is decreasing because, as the number of selected sensor nodes grows larger, the energy cost in each information acquisition is also increased. The performance with the optimized sensor nodes selection scheme is better than the conventional sensor nodes selection scheme. By using the same energy cost, the proposed scheme can achieve more information acquisition times compared with the traditional scheme. Therefore, it can prolong the network lifetime.

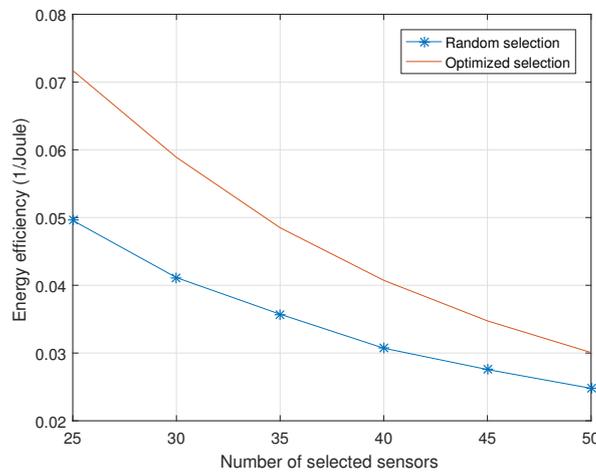


Figure 5. Energy efficiency versus different number of selected sensor nodes.

5. Conclusions

In this work, an optimized sensor node selection scheme for compressed information acquisition is provided for improving the accuracy of reconstruction and prolonging the network lifetime. The signal reconstruction is achieved based on Bayesian estimation theory. The MSE of estimation can be obtained by the given error covariance matrix. Then, the proposed optimization problem aims at minimizing it by taking the selection matrix as the variable. By relaxing the non-convex constraint, the proposed optimization problem can be solved effectively. To increase the sustainability of the network, the scheme selects the nodes with more energy by adding a bound constraint in the

optimization problem. Compared with the conventional node selection scheme, simulation results demonstrate that the proposed scheme can achieve more accurate reconstruction and improve the network lifetime significantly.

Author Contributions: R.W. and G.L. conceived and designed the experiments, G.L. and W.K. analyzed the data, B.L. and R.M. performed the experiments, R.W. wrote the paper, C.Z. enhanced the paper.

Funding: This work is supported by National Natural Science Foundation of China (Nos. 61501139, 61371100, 61401118).

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

NP	Non-deterministic polynomial
CSI	Channel state information
MSE	Mean square error
EM	Expectation maximization
OMP	Orthogonal matching pursuit
OOMP	OMP algorithm with optimized sensor nodes selection scheme
BCS	Bayesian compression sensing estimation
OBCS	Bayesian compression sensing estimation with optimized sensor nodes selection scheme

Appendix A

To achieve the final sparsity prior, hierarchical priors about the hyperparameters α , σ^2 and the sparse vector \mathbf{s} are given in this section. The hyperprior of α is defined as a Gamma prior distribution:

$$\begin{aligned} p(\alpha) &= \prod_{i=1}^N \Gamma(\alpha_i | a, b) \\ &= \prod_{i=1}^N \int_0^{+\infty} x^{\alpha-1} e^{-x} dx b^a \alpha^{a-1} e^{-b\alpha}. \end{aligned} \quad (\text{A1})$$

It can be seen that the prior information is not sufficient because the prior distribution of the variance σ^2 is unknown. The hyperprior of σ^{-2} as a Gamma prior distribution is defined as follows:

$$p(\sigma^{-2}) = \prod_{i=1}^N \Gamma(\sigma^{-2} | c, d). \quad (\text{A2})$$

If $a = b = c = d = 0$ holds, the hyperprior distribution becomes uniform distribution. In this paper, this assumption is kept. It is convenient to discuss the estimation of hyperparameters.

The posterior distribution of the unknown parameters \mathbf{s} , α , and σ^2 is given by Bayes's rules

$$p(\mathbf{s}, \alpha, \sigma^2 | \mathbf{y}) = \frac{p(\mathbf{y} | \mathbf{s}, \alpha, \sigma^2) p(\mathbf{s}, \alpha, \sigma^2)}{p(\mathbf{y})}. \quad (\text{A3})$$

Considering that it is difficult to calculate $p(\mathbf{y})$, which is the integral of $p(\mathbf{y} | \mathbf{s}, \alpha, \sigma^2) p(\mathbf{s}, \alpha, \sigma^2)$, we decompose the posterior $p(\mathbf{s}, \alpha, \sigma^2 | \mathbf{y})$ as follows instead of directly calculating it:

$$p(\mathbf{s}, \alpha, \sigma^2 | \mathbf{y}) = p(\mathbf{s} | \alpha, \sigma^2, \mathbf{y}) p(\alpha, \sigma^2 | \mathbf{y}). \quad (\text{A4})$$

Then, the posterior of the sparse vector \mathbf{s} can be obtained:

$$p(\mathbf{s} | \alpha, \sigma^2, \mathbf{y}) = \frac{p(\mathbf{y} | \mathbf{s}, \sigma^2) p(\mathbf{s} | \alpha)}{p(\mathbf{y} | \alpha, \sigma^2)}. \quad (\text{A5})$$

References

1. Zhu, C.S.; Shu, L.; Leung, V.C.M.; Guo, S.; Zhang, Y.; Yang, L.T. Secure multimedia big data in trust-assisted sensor-cloud for smart city. *IEEE Commun. Mag.* **2017**, *55*, 24–30. [[CrossRef](#)]
2. Zhu, C.S.; Wang, H.; Liu, X.L.; Shu, L.; Yang, L.T.; Leung, V.C.M. A novel sensory data processing framework to integrate sensor networks with mobile cloud. *IEEE Syst. J.* **2016**, *10*, 1125–1136. [[CrossRef](#)]
3. Han, Y.F.; Zhang, J.C.; Sun, D.J. Error control and adjustment method for underwater wireless sensor network localization. *Appl. Acoust.* **2018**, *130*, 293–299. [[CrossRef](#)]
4. Zhang, J.R.; Xia, Z.X.; Huang, L.Y.; Ma, L.K. Experimental and numerical parametric studies on two-Phase underwater ramjet. *J. Propuls. Power* **2018**, *34*, 161–169. [[CrossRef](#)]
5. Sun, F.F.; Liu, W.; Dong, Z.X.; Deng, Y.L. Underwater superoleophobicity cellulose nanofibril aerogel through regioselective sulfonation for oil/water separation. *Chem. Eng. J.* **2017**, *330*, 774–782. [[CrossRef](#)]
6. Wang, C.; Yu, H.Y.; Zhu, Y.J.; Wang, T. Blind detection for SPAD-based underwater VLC system under P-G mixed noise model. *IEEE Commun. Lett.* **2017**, *21*, 2602–2605. [[CrossRef](#)]
7. Han, G.J.; Jiang, J.F.; Bao, N.; Wan, L.T.; Guizani, M. Routing protocols for underwater wireless sensor networks. *IEEE Commun. Mag.* **2015**, *53*, 72–78. [[CrossRef](#)]
8. Han, G.J.; Jiang, J.F.; Sun, N.; Shu, L. Secure communication for underwater acoustic sensor networks. *IEEE Commun. Mag.* **2015**, *53*, 54–60. [[CrossRef](#)]
9. Zhu, C.S.; Li, X.H.; Leung, V.C.M.; Yang, L.T. Towards pricing for sensor-cloud. *IEEE Trans. Cloud Comput.* **2017**. [[CrossRef](#)]
10. Zhu, C.S.; Leung, V.C.M.; Wang, K.; Yang, L.T.; Zhang, Y. Multi-method data delivery for green sensor-cloud. *IEEE Commun. Mag.* **2017**, *55*, 176–182. [[CrossRef](#)]
11. Li, S.C.; Xu, L.D.; Wang, X.H. Compressed sensing signal and data acquisition in wireless sensor networks and internet of things. *IEEE Trans. Ind. Inform.* **2013**, *9*, 2177–2186. [[CrossRef](#)]
12. Do, T.T.; Gan, L.; Nguyen, N.H.; Tran, T.D. Fast and efficient compressive sensing using structurally random matrices. *IEEE Trans. Signal Process.* **2012**, *60*, 139–154. [[CrossRef](#)]
13. Duarte, M.F.; Eldar, Y.C. Structured compressed sensing: From theory to applications. *IEEE Trans. Signal Process.* **2011**, *59*, 4053–4085. [[CrossRef](#)]
14. Bayati, M.; Montanari, A. The dynamics of message passing on dense graphs, with applications to compressed sensing. *IEEE Trans. Inf. Theory* **2011**, *57*, 764–785. [[CrossRef](#)]
15. Adcock, B.; Hansen, A.C. Generalized sampling and infinite-dimensional compressed sensing. *Found. Comput. Math.* **2016**, *16*, 1263–1323. [[CrossRef](#)]
16. Bajwa, W.U.; Haupt, J.; Sayeed, A.M.; Nowak, R. Compressed channel sensing: A new approach to estimating sparse multipath channels. *Proc. IEEE* **2010**, *98*, 1058–1076. [[CrossRef](#)]
17. Haupt, J.; Bajwa, W.U.; Raz, G.; Nowak, R. Toeplitz compressed sensing matrices with applications to sparse channel estimation. *IEEE Trans. Inf. Theory* **2010**, *56*, 5862–5875. [[CrossRef](#)]
18. Carlin, M.; Rocca, P.; Oliveri, G.; Viani, F.; Massa, A. Directions-of-arrival estimation through Bayesian compressive sensing strategies. *IEEE Trans. Antennas Propag.* **2013**, *61*, 3828–3838. [[CrossRef](#)]
19. Lustig, M.; Donoho, D.; Pauly, J.M. Sparse MRI: The application of compressed sensing for rapid MR imaging. *Magn. Reson. Med.* **2007**, *58*, 1182–1195. [[CrossRef](#)] [[PubMed](#)]
20. Yang, L.; Song, K.; Siu, Y.M. Iterative clipping noise recovery of OFDM signals based on compressed sensing. *IEEE Trans. Broadcast.* **2017**, *63*, 706–713. [[CrossRef](#)]
21. Bevacqua, M.T.; Crocco, L.; Di Donato, L.; Isernia, T. Non-linear inverse scattering via sparsity regularized contrast source inversion. *IEEE Trans. Comput. Imaging* **2017**, *3*, 296–304. [[CrossRef](#)]
22. Bevacqua, M.T.; Isernia, T. Shape reconstruction via equivalence principles, constrained inverse source problems and sparsity promotion. *Prog. Electromagn. Res.* **2017**, *158*, 37–48. [[CrossRef](#)]
23. Bevacqua, M.T.; Scapatucci, R. A compressive sensing approach for 3D breast cancer microwave imaging with magnetic nanoparticles as contrast agent. *IEEE Trans. Med. Imaging* **2016**, *35*, 665–673. [[CrossRef](#)] [[PubMed](#)]
24. Sung, C.K.; de Hoog, F.; Chen, Z.; Cheng, P.; Popescu, D.C. Interference mitigation based on Bayesian compressive sensing for wireless localization systems in unlicensed band. *IEEE Trans. Veh. Technol.* **2017**, *66*, 7038–7049. [[CrossRef](#)]

25. Gu, X.J.; Zhou, P.; Gu, X.S. Bayesian compressive sensing for thermal imagery using Gaussian-Jeffreys prior. *Infrared Phys. Technol.* **2017**, *83*, 51–61. [[CrossRef](#)]
26. Hawes, M.; Mihaylova, L.; Septier, F.; Godsill, S. Bayesian compressive sensing approaches for direction of arrival estimation with mutual coupling effects. *IEEE Trans. Antennas Propag.* **2017**, *65*, 1357–1368. [[CrossRef](#)]
27. Ding, W.B.; Lu, Y.; Yang, F.; Dai, W.; Li, P.; Liu, S.C.; Song, J. Spectrally efficient CSI acquisition for power line communications: A Bayesian compressive sensing perspective. *IEEE J. Sel. Areas Commun.* **2016**, *34*, 2022–2032. [[CrossRef](#)]
28. Tipping, M.E. Sparse Bayesian learning and the relevance vector machine. *J. Mach. Learn. Res.* **2001**, *1*, 211–244.
29. Ozgor, M.; Erkucuk, S.; Cirpan, H.A. Bayesian compressive sensing for ultra-wideband channel estimation: Algorithm and performance analysis. *Telecommun. Syst.* **2015**, *59*, 417–427. [[CrossRef](#)]
30. Yu, F.Y.; Li, D.Z.; Guo, Q.; Wang, Z.Y.; Xiang, W. Block-FFT based OMP for compressed channel estimation in underwater acoustic communications. *IEEE Commun. Lett.* **2015**, *19*, 1937–1940. [[CrossRef](#)]
31. Fazel, F.; Fazel, M.; Stojanovic, M. Random access compressed sensing for energy-efficient underwater sensor networks. *IEEE J. Sel. Areas Commun.* **2011**, *29*, 1660–1670. [[CrossRef](#)]
32. Lin, H.Z.; Wei, W.; Zhao, P.; Ma, X.Q.; Zhang, R.; Liu, W.P.; Deng, T.P.; Peng, K. Energy-efficient compressed data aggregation in underwater acoustic sensor networks. *Wirel. Netw.* **2016**, *22*, 1985–1997. [[CrossRef](#)]
33. Yan, H.C.; Xu, J.; Long, T.; Zhang, X.D. Underwater acoustic matched field imaging based on compressed sensing. *Sensors* **2015**, *15*, 25577–25591. [[CrossRef](#)] [[PubMed](#)]
34. Cheung, B.C.; Carriveau, R.; Ting, D.S.K. Parameters affecting scalable underwater compressed air energy storage. *Appl. Energy* **2014**, *134*, 239–247. [[CrossRef](#)]
35. Huang, K.D.; Guo, Y.; Guo, X.M.; Wang, G.L. Heterogeneous Bayesian compressive sensing for sparse signal recovery. *IET Signal Process.* **2014**, *8*, 1009–1017. [[CrossRef](#)]
36. Wimalajeewa, T.; Varshney, P.K. OMP based joint sparsity pattern recovery under communication constraints. *IEEE Trans. Signal Process.* **2014**, *62*, 5059–5072. [[CrossRef](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).